

Human N-cadherin: nucleotide and deduced amino acid sequence

Robert A. Reid and John J. Hemperly

Becton Dickinson and Company Research Center, PO Box 12016, Research Triangle Park, NC 27709, USA

Submitted August 14, 1990

EMBI accession no X54315

N-cadherin, a member of the cadherin gene family originally detected in nervous tissue, is a transmembrane protein which mediates calcium-dependent, homophilic, cell-cell adhesion (1). We report here the entire protein coding and additional 5' and 3' untranslated nucleotide sequences of human N-cadherin as well as the deduced amino acid sequence. An N-cadherin probe was generated by PCR amplification of human brain polyA+ RNA using oligonucleotide primers designed from the chicken N-cadherin sequence (2). This probe was used to screen a Kelly neuroblastoma library (Clontech, Palo Alto, CA) from which fifteen independent clones, one being full length, were obtained. The reported sequence was determined on both strands by partial sequencing of multiple cDNAs and was defined as N-cadherin by comparison to N-cadherin isolated from chicken and mouse (3). Nucleotide sequence identities between human/mouse and human/chicken are 96% and 85%, respectively. Human and

mouse sequences lack two regions of five and one amino acid each, which are present in chicken. The five amino acid deletion exists within the putative signal sequence between amino acid positions -146 and -147. The single amino acid deletion is present within the extracellular domain between amino acid positions 373 and 374. The signal sequence (solid squares), membrane-spanning region (solid circles), cadherin extracellular repeats (overlines), amino terminus of the mature polypeptide (arrow) and potential N-glycosylation sites (circles) are indicated. The nucleotide and amino acid positions are numbered on the right and left, respectively.

REFERENCES

1. Takeichi,M. (1988) *Development* **102**, 639–655.
 2. Hatta,K. et al. (1988) *J. Cell Biol.* **106**, 873–881.
 3. Miyatani,S. et al. (1989) *Science* **245**, 631–635.

ACCCTGGCCGCTGTGGTGTGCTGCCGCTGCCCTCCTCCCTCCGCCGCCGCCGCCGCCCTCTCGCTCGGCCCTCTCCGCCCTCCATGTGCGGAGATAGCGGAGAC
 -150 M C R E I B A G C 1200
 CGCTGGGACCTCTGTGCGCTGCCCTGCTGGCTGGCCCTGCTCAGGGCTGTAGAGGCTTCTGGTAAATGCCATTATGCCAAGACTGGATTCTCGAACATTTACAGTCAGTC
 -153 A B L R T B L L P S L L B L Q A S V E A S G E I A L C K T G F P E D V Y S A V L S 2400
 AGGATGGCATGAAGGACAGCCTCTCTCAATGTGAAGGTTAGCAATGCCAATGGAAAAAAAGAACATTAAGGAGACTGGAGCTGACGCCAGCTGCCAGATTTCAGGATGC
 -113 K D V H E G C P L L N V K F S N C N G K R K V Y E S S E P A D F K V D E D G M 3600
 TGATGCCGTGAGAACCTTCCACTCTCTTGACCATGCCAAGTCTCTGATATATGCCAACGACAAGAACAGGCCAGAAAAGTGGCAACTGGCAAGTAAATTGAGCCTGAAGCAACCT
 -73 V Y A V R S F P L S S E H A K F L I Y A R D K E T R E K W Q V A V K L S L K P T 4800
 TAATCTGAGGAGTCAGTGAAGGAGTCAGCACAGAACAGAATAGTGTCTCCAAAGCACATTGCAAGCACAGTGGCACCTACAAAGCAGAACAGAGACACTGGCTCATGCCCTCAATCA
 -33 L T E E S V K E S A E V E E I V F P R F S K H G L B R Q K R D -W-V-I-P- 6000
 ACTTGCCAGAAAATCTCAGGGACCTTCTCAAGACCTTGCTCAGCATCAGGCTCTACAGATAAACCTTCCGCTGCCGACTCTGTAACTGCCAGGAGCTGCCAGGCCCTCCAA
 8 N L P E N S R G P F P Q E L V R I R S D R D K (N) L S L R Y T V T G P G A D P P 7200
 CTGGTAGTCTTCATTATCAACCCCATTCGGGTGAGCTGCCGACCATAGGCCGGTTTCTATTGAGGCCATGCACTAGATATTAAATGGAAAATCAAG
 48 T G I F I N P I S G Q L S V T K P L D R E Q I A R F H L R A H A V D I N G N Q 8400
 TGGGAACCCCCATTGACATTGTCAATGTTATTGACATGAATCACAAAGACCTGAGTCTTACACCAAGGGTTGGAAATGGGACAGTCTCTGAGGCCATCAAGCCTGGAAATGTGA
 86 V E N P I D L V I N V D O M N D N R P E F L H B V W (N) G T V P E G S K P G T Y V 9600
 TGACCGTAACGCAATTGATGTGAGCATCCAACTGCCCTCAATGGATGTGAGGTAACATGGCTGTCTCAGGCTCAACGCCCTTACCCCAACATGTTAACATCAACAAATGAGA
 128 M T V T I D A D D P P N A L G N M L R Y R I V S Q R A P S T P S P N M F T I N (N) E 10800
 CTGGTAGCATCATCACAGTGGCAGCTGACTGATCACAGAACAGCTGCAACAGTATACGTTAAATTAACGGTACACAGCATGGAAAGATCCCCACATATGCCCTTCAACACAGGCCA
 168 T G D I I T V A A G L D R E N V Q B Y T L I I I G A T D M E G I P T Y G C L S N T A 12000
 CGCCGCTCATCACAGTGGCAGCTGACTGATCACAGAACAGCTGCAACAGTATACGTTAAATTAACGGTACACAGCATGGAAAGATCCCCACATATGCCCTTCAACACAGGCCA
 206 T A V I T V D T V N D N P P E F T A T M T F Y G E V P E N R V D I I V A (N) L T V T 13200
 ATAAGGATCAACCCCCATACACCGCTGGAACCGCAGTGACAGAATCACTGGCGGAGATCTACTGGACGGTTGCCATCCAGGCCACCCAAAGCAACGCCGGTTAGTCAACCGTGG
 248 D K D Q P H T P A W N A V Y R I S G C D P T R G F R A I S T D P N S N D G L V T V 14400
 TCAAAACATCGACTTGTGAAACATAGGATGTTCTCTTACTGTGAGGAGCAAGAACATTCAGGCTCAACGCCCTTACGGCAACGGGAAATTGCCACCCSCTCACTGCAACCTGGCTCTGTTA
 286 V K P I D F E T N R M F V L T V A A E N O V P L A K G I Q H P P R S T A T V S V 15600
 CAGTTATTGAGCTAAATGAAAACCTTATTGCCCCCACTTCAAGATCATTGCCAACGAAAGGGCTTACCTGGACATTGAGCAACATTCACTGCAAGCCCAGATCGAT
 328 T V I D V N E N P Y F A P C P N P K I I R G E E G L H A G T M L T T F T A Q R D P D R 16800
 ATATCGAGCAAATATTAGATACAATTATCTGATCTGCCAATGGCTAAAAGATAGTCTGGTAATGGACACAATACTACATGTTCTGTTGGACCCAGAACATCACCAGATGTGA
 368 Y M Q G N O R I Y T K L S D P A N W L K I D P V N G R I T T I A V L D R E S P N V 18000
 AAAACAAATATATAATGACTTCTCTTGCTTCAACATGCCAATTCTCTATGAGTGGCAACGCCGCTGCCACATCTATTACTGATATTAAATGACAATGCCCTCAAGTGTAC
 408 K N N I Y (N) A T F L A S D N G I P P M S G T G L T L L D I N D N A P Q V L 19200
 CTCAGAGGCCAGACATCTGGCAACTCCAGGCCCAATTCAATTATACAGCAGCTGATTATGACATGGCTTCTTCTTCTTCTCCAGTG
 448 P Q E A T C E T C P D P N S I (N) I T A L D Y D I D P N A G F A F D L P L S P V 20400
 CTATTAAGGAAATTGGACCATCACTGGCTTAATGCTGATTTCAGCTTAATTAAAGATAAAATTCTGAAACTGCTGTATCTGAACTTCCCATCATTAATCACAGGTTCTGGGAT
 488 T I K R (N) W T I T R L N G D F A Q G L N L K I K F L E A G I Y E V P I I T D S C 21600
 ATCTCCCAAATCAATTTCATCTGCCGTGGGGTGGCAACTGACTTCAGGCCAGACTGAGCATGGCTGGGGCTGGGGCTGGCAACGGCTGCCATCTGGGAT
 528 N P P K S (N) I S I L R V R V C Q C D F N G D C T D V D R I O V C G A O G C O L @ C T O I @ I @ I @ 22800
 CCATCTGCTGCTGACATCATCTGCCCTTATCTCTGGCTGATTTGGTGTATGGATGAAAGGCCGGGATAAAGAACGCCAGGCCAAACAACCTTTAATGATGCCAGAACATGATGTAA 2400
 568 A @ I O L @ C O I @ I @ I O L @ I O L @ V O L @ M O F V @ V @ W @ M @ R @ K @ R @ D @ K @ R @ Q @ K @ L @ I D P E D D V 25200
 GAGATAATTAAATGATCAACAGGTGGAGGAGAACAGAACGACCTGAGCTGACTGATCTGGCCAGCTGGAGCTGACACTGGCTGGCCCTGATGCCCATCAAGGCCATGGGAATCC
 608 R D N I L K Y D E E G G E D Q D Y D L S Q L R Q P D T V E P D A I K P V G I 26400
 GCAACATGGATGAAAGGCCATCACTGGCTGAGGCCAGTATCCGGTCCGGATCTGAGGCCACACCTGGGACTCTGATTTCAATGAGGCCCTAAAGGCCCTCAACATGCC
 648 R R M D E R P I H A E P Y P V R S A P H P G D I G D F I N E G L K A A D D N D 27600
 CCACACTCCACCATATGACTCTCTGTAGTTTACTATGAGGECAGTGGCTCCACTGCTGGCTCTTGAGCTCCCTTAATTCCTCAAGTAGTGGTGTAGGAGGACTATGATTAC
 688 P T A P P Y P D S L L V F D Y E G S G S T A G S L S L S S S S G C E Q D Y D Y 28800
 TGAACGACTGGGGCAGGGGCACTGGGACTGAGCATGTGATGGTGGAGGTGACTGACTGAACTTCAAGTAGTGGTGTAGGAGGACTATGATTAC
 728 L N D W G P R F K K L A D M Y G G C D D * 2903